

TITLE **AP20 Rec'd PCT/PTO 19 JUL 2006****ALTERING ROOT STRUCTURE DURING PLANT DEVELOPMENT****FIELD OF THE INVENTION**

5       The field of invention relates to plant breeding and genetics and, in particular, relates to recombinant constructs useful for altering root structure during plant development.

**BACKGROUND OF THE INVENTION**

10       Relatively little is known about the genetic regulation of plant root development and function. Elucidation of the genetic regulation is important because roots serve important functions such as acquisition of water and nutrients and the anchorage of the plants in the soil.

15       The mutation of the *RTCS* (rootless for crown and seminal roots) gene was first described by Hetz et al. (1996) *Plant J.* 10(5):845-857. Two maize *rtcs* mutants, *rtcs1* and *rtcs2*, were shown to have reduced resistance to root lodging. Both mutants were found to be deficient in formation of both crown and seminal lateral roots, which appear to be suppressed at a very early stage in mutant embryos. In addition, brace root formation, which occurs at a later stage of development in wild type-plants, was also found to be lacking in the two *rtcs* mutants. Microscopic  
20       analysis further showed that root primordia formation was absent in mutant plants.

25       The mutation was genetically mapped to the short arm of chromosome 1, with apparently no phenotypic differences between the two mutants. Genetic analysis of the two *rtcs* mutations indicate that they are both inherited as monogenic recessive traits. More extended mapping analysis and a narrowing of the location of the mutant locus on chromosome 1 was performed as described in Krebs et al, (1999) *MNL* 73:33.

30       Despite the extensive genetic and morphological characterization of the *rtcs* mutants, there has been no molecular analysis of the nucleic acid encoding the protein associated with the *rtcs* phenotype. Indeed, the identity of the protein encoded by *RTCS* has not been reported. A hypothetical protein sequence of 287 amino acids from rice is disclosed in the NCBI Database at Accession No. AAN87738.1 (GI No. 27261472).

35       It has been reported that the Lateral Organ Boundary (LOB) gene in *Arabidopsis* has a potential role in lateral organ development. See Shuai et al., (2002), *Plant Phys.* 129, 747-761. Shuai et al. found LOB gene expression at the base of lateral organs in the shoots and roots of *Arabidopsis*. Moreover, 23 members of the LOB domain family (LBD) of genes were found to exhibit expression patterns in the root tissues of *Arabidopsis*.

### SUMMARY OF THE INVENTION

The present invention includes isolated polynucleotides encoding a polypeptide required for proper root formation, wherein the amino acid sequence of the polypeptide and the amino acid sequence of SEQ ID NO: 6, 8, 30, or 38 have at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% identity, or the complement of the nucleotide sequence, wherein the complement and the nucleotide sequence contain the same number of nucleotides and are 100% complementary. The polypeptide preferably comprises the amino acid sequence of SEQ ID NO: 6, 8, 30, or 38. The nucleotide sequence preferably comprises the nucleotide sequence of SEQ ID NO: 5, 7, 29, or 37.

In a first embodiment, the present invention includes an isolated polynucleotide comprising: (a) a nucleotide sequence encoding a polypeptide required for proper root formation, wherein the polypeptide has an amino acid sequence of at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% sequence identity based on the Clustal V method of alignment when compared to a polypeptide SEQ ID NO: 6, 8, 30, or 38, or (b) a complement of the nucleotide sequence, wherein the complement and the nucleotide sequence contain the same number of nucleotides and are 100% complementary.

In a second embodiment, this invention includes isolated polynucleotide sequences or complements comprising at least two motifs corresponding substantially to any of the amino acid sequences set forth in SEQ ID NOs: 9, 10, 11, 12 or 13, wherein said motif is substantially a conserved subsequence.

In a third embodiment, this invention includes a functionally equivalent subfragment of an isolated polynucleotide (or complement) of the present invention, wherein the subfragment is useful in antisense inhibition or co-suppression of a protein altering root structure in a transformed plant.

In a fourth embodiment, this invention includes an isolated nucleic acid fragment comprising a promoter wherein said promoter consists essentially of the nucleotide sequence set forth in SEQ ID NO: 1, 2, 3 or 4 or said promoter consists essentially of a fragment or subfragment that is substantially similar and functionally equivalent to the nucleotide sequence set forth in SEQ ID NO's: 1, 2, 3 or 4.

In a fifth embodiment, this invention includes recombinant DNA constructs comprising any of the foregoing nucleic acid fragments or complements or functionally equivalent subfragments, operably linked to at least one regulatory sequence. Also included are plants comprising such recombinant DNA constructs in their genome, plant tissue or cells obtained from such plants, and seeds obtained from these plants.

In a sixth embodiment, this invention includes a method of altering root structure during plant development in plants which comprises:

- (a) transforming a plant with a recombinant DNA construct of the invention;
- (b) growing the transformed plant under conditions suitable for the expression of the recombinant DNA construct; and
- (c) selecting those transformed plants with suppresses root formation.

In a seventh embodiment, this invention includes a method to isolate nucleic acid fragments encoding polypeptides associated with altering root structure during plant development which comprises:

- (a) comparing SEQ ID NO's: 6, 8, 30, or 38 with other polypeptide sequences associated with altering root structure during plant development;
- (b) identifying the conserved sequences(s) or 4 or more amino acids obtained in step (a);
- (c) making region-specific nucleotide probe(s) or oligomer(s) based on the conserved sequences identified in step (b); and
- (d) using the nucleotide probe(s) or oligomer(s) of step (c) to isolate sequences associated with altering root structure during plant development by sequence dependent protocols.

In an eighth embodiment, this invention also includes a method of mapping genetic variations related to altering root structure during plant development comprising:

- (a) crossing two plant varieties; and
  - (b) evaluating genetic variations with respect to:
    - (i) a nucleic acid sequence selected from the group consisting of SEQ ID NO: 5, 7, 29, or 37; or
    - (ii) a nucleic acid sequence encoding a polypeptide selected from the group consisting of SEQ ID NO: 6, 8, 30, or 38;
- in progeny plants resulting from the cross of step (a), wherein the evaluation is made using a method selected from the group consisting of: RFLP analysis, SNP analysis, and PCR-based analysis.

In a ninth embodiment, this invention includes a method of molecular breeding to alter root structure during plant development, the method comprising:

- (a) crossing two plant varieties; and
- (b) evaluating genetic variations with respect to:
  - (i) a nucleic acid sequence selected from the group consisting of SEQ ID NO: 5, 7, 29, or 37; or
  - (ii) a nucleic acid sequence encoding a polypeptide selected from the group consisting of SEQ ID NO: 6, 8, 30, or 38

in progeny plants resulting from the cross of step (a), wherein the evaluation is made using a method selected from the group consisting of: RFLP analysis, SNP analysis, and PCR-based analysis.

## BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCE LISTINGS

The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing, which form a part of this application.

Figure 1 shows an alignment of the RTCS polypeptide sequence from maize inbred line Mo17 (SEQ ID NO:6), the deduced rice polypeptide sequence (SEQ ID NO:8), the RTCS polypeptide sequence from maize inbred line B73 (SEQ ID NO:30), a hypothetical protein sequence from rice (SEQ ID NO: 26, GI No. 27261472), and a putative LOB domain protein sequence from *Arabidopsis* (SEQ ID NO: 27, GI No. 22331847).

Figure 2 shows vector PHP24451.

Figure 3 shows vector PHP24452.

Figure 4 shows the rtcs allele described in Example 10.

SEQ ID NO:1 represents the 6101 bp of the maize genomic sequence containing the ORF (Nucleotides 2779-3610) of the RTCS gene, which is interrupted by an intron extending from Nucleotides 3060-3159.

SEQ ID NO:2 is 2778 bp, extending from Nucleotides 1-2778 of the putative promoter of the maize RTCS genomic sequence shown in SEQ ID NO:1.

SEQ ID NO:3 is 2000 bp, extending from Nucleotides 779 –2778 of the putative promoter of the maize RTCS genomic sequence shown in SEQ ID NO:1.

SEQ ID NO:4 is 1000 bp, extending from Nucleotide 1779-2778 of the putative promoter of the maize RTCS genomic sequence shown in SEQ ID NO:1.

SEQ ID NO:5 is the nucleotide sequence of the ORF of SEQ ID NO:1 minus an intron sequence (i.e., SEQ ID NO:5 is nucleotides 2779-3059 and 3160-3610 of SEQ ID NO:1).

SEQ ID NO:6 is the amino acid sequence encoded by SEQ ID NO:5.

SEQ ID NO:7 is the deduced nucleotide sequence for the coding sequence of the rice RTCS gene.

SEQ ID NO:8 is the deduced amino acid sequence encoded by SEQ ID NO:7.

SEQ ID NO:9 is a conserved sequence motif associated with nucleotide sequences included in the present invention.

SEQ ID NO:10 is a conserved sequence motif associated with nucleotide sequences included in the present invention.

SEQ ID NO:11 is a conserved sequence motif associated with nucleotide sequences included in the present invention.

SEQ ID NO:12 is a conserved sequence motif associated with nucleotide sequences included in the present invention.

5 SEQ ID NO:13 is a conserved sequence motif associated with nucleotide sequences included in the present invention.

SEQ ID NO:14 is the forward primer for SSR marker BNLG1014 used in Example 1.

10 SEQ ID NO:15 is the reverse primer for SSR marker BNLG 1014 used in Example 1.

SEQ ID NO:16 is the forward primer for SSR marker BNLG 1429 used in Example 1.

SEQ ID NO:17 is the reverse primer for SSR marker BNLG 1429 used in Example 1.

15 SEQ ID NO:18 is the forward primer for SSR marker UMC1685 used in Example 1.

SEQ ID NO:19 is the reverse primer for SSR marker UMC 1685 used in Example 1.

20 SEQ ID NO:20 is the forward primer for SSR marker UMC1660 used in Example 1.

SEQ ID NO:21 is the reverse primer for SSR marker UMC 1660 used in Example 1.

SEQ ID NO:22 is the forward primer for Cap marker b104.i24 used in Example 1.

25 SEQ ID NO:23 is the reverse primer for Cap marker b104.i24 used in Example 1.

SEQ ID NO:24 is the forward primer of Cap marker b74.m9 used in Example 1.

30 SEQ ID NO:25 is the reverse primer of Cap marker b74.m9 used in Example 1.

SEQ ID NO:26 shows the amino acid sequence for the rice hypothetical protein gi: 27261472.

SEQ ID NO:27 shows the amino acid sequence for the Arabidopsis putative LOB domain protein gi: 22331847.

35 SEQ ID NO:28 represents the 3286 bp of the maize genomic sequence containing the ORF (Nucleotides 1200 - 2030) of the RTCS gene, which is interrupted by an intron extending from Nucleotides 1482-1577.

SEQ ID NO:29 is the nucleotide sequence of the ORF of SEQ ID NO:28 minus an intron sequence (i.e., SEQ ID NO:29 is nucleotides 1200-1481 and 1578-2030 of SEQ ID NO:28).

SEQ ID NO:30 is the amino acid sequence encoded by SEQ ID NO:29.

SEQ ID NO:31 is the forward primer used in Example 4.

SEQ ID NO:32 is the reverse primer used in Example 4.

SEQ ID NO:33 is the forward primer used in Example 6.

SEQ ID NO:34 is the reverse primer used in Example 6.

SEQ ID NO:35 is the forward primer used in Example 9.

SEQ ID NO:36 is the reverse primer used in Example 9.

SEQ ID NO:37 is the RTCS-like cDNA obtained from the experiment described in Example 9.

SEQ ID NO:38 is the amino acid sequence encoded by SEQ ID NO: 37.

SEQ ID NO:39 is the genomic sequence containing the RTCS-like gene .

SEQ ID NO:40 is the forward primer used in Example 10.

SEQ ID NO:41 is the reverse primer used in Example 10.

SEQ ID NO:42 is the 3864 bp nucleotide sequence containing the 1536bp of the rtcs mutant gene described in Example 10.

The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res.* 13:3021-3030 (1985) and in the *Biochemical J.* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

#### DETAILED DESCRIPTION OF THE INVENTION

The disclosure of each reference set forth herein is incorporated herein by reference in its entirety.

As used herein, an "isolated nucleic acid fragment" is used interchangeably with "isolated polynucleotide" and is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or

DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

5       The term "isolated" refers to materials, such as nucleic acid molecules and/or proteins, which are substantially free or otherwise removed from components that normally accompany or interact with the materials in a naturally occurring environment. Isolated polynucleotides may be purified from a host cell in which they naturally occur. Conventional nucleic acid purification methods known to  
10 skilled artisans may be used to obtain isolated polynucleotides. The term also embraces recombinant polynucleotides and chemically synthesized polynucleotides.

      The terms "subfragment that is functionally equivalent" and "functionally equivalent subfragment" are used interchangeably herein. These terms refer to a  
15 portion or subsequence of an isolated nucleic acid fragment in which the ability to alter gene expression or produce a certain phenotype is retained whether or not the portion or subsequence encodes an active enzyme or functional protein (for example, the portion or subsequence may be a portion of coding and/or non-coding regions and need not encode an active enzyme or functional protein. For example,  
20 the fragment or subfragment can be used in the design of recombinant DNA constructs to produce the desired phenotype in a transformed plant. Recombinant DNA constructs can be designed for use in co-suppression or antisense by linking a nucleic acid fragment or subfragment thereof, whether or not it encodes an active enzyme or functional protein, in the appropriate orientation relative to a plant  
25 promoter sequence.

      The terms "homology", "homologous", "substantially similar" and "corresponding substantially" are used interchangeably herein. They refer to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate gene expression or produce a  
30 certain phenotype. These terms also refer to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially alter the functional properties of the resulting nucleic acid fragment relative to the initial, unmodified fragment. It is therefore understood, as those skilled in the art will appreciate, that the invention  
35 encompasses more than the specific exemplary sequences.

      Moreover, the skilled artisan recognizes that substantially similar nucleic acid sequences encompassed by this invention are also defined by their ability to hybridize, under moderately stringent conditions (for example, 1 X SSC, 0.1% SDS,

60°C) with the sequences exemplified herein, or to any portion of the nucleotide sequences reported herein and which are functionally equivalent to the gene or the promoter of the invention. Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly  
5 related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. One set of preferred conditions involves a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with  
10 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions involves the use of higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions involves the use of two final washes in 0.1X SSC, 0.1% SDS at 65°C.

15 With respect to the degree of substantial similarity between the target (endogenous) mRNA and the RNA region in the construct having homology to the target mRNA, such sequences should be at least 25 nucleotides in length, preferably at least 50 nucleotides in length, more preferably at least 100 nucleotides in length, again more preferably at least 200 nucleotides in length, and most  
20 preferably at least 300 nucleotides in length; and should be at least 80% identical, preferably at least 85% identical, more preferably at least 90% identical, and most preferably at least 95% identical.

Sequence alignments and percent similarity calculations may be determined using a variety of comparison methods designed to detect homologous sequences  
25 including, but not limited to, the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences are performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments and  
30 calculation of percent identity of protein sequences using the Clustal method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4.

35 "Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Recombinant DNA construct" refers to a combination of nucleic acid fragments that are not normally found



together in nature. Accordingly, a recombinant DNA construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that normally found in nature. A  
5 "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or recombinant DNA constructs. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

10 "Coding sequence" refers to a DNA sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence.

15 Regulatory sequences may include, but are not limited to, promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

"Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as  
20 enhancers. Accordingly, an "enhancer" is a DNA sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoter sequences can also be located within the transcribed portions of genes, and/or downstream of the transcribed sequences. Promoters may be derived in their  
25 entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of an isolated nucleic acid fragment in different tissues or cell types, or at different stages of development, or in response to different environmental  
30 conditions. Promoters which cause an isolated nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg, (1989) *Biochemistry of Plants* 15:1-82.

35 It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. As used herein, "substantially similar and functionally equivalent subfragment of a promoter" refers to a portion or

subsequence of a promoter sequence which is capable of controlling the expression of a coding sequence or functional RNA.

Specific examples of promoters that may be useful in expressing the nucleic acid fragments of the invention include, but are not limited to, the promoters disclosed in this application (SEQ ID NO's: 1,2,3 and 4).

An "intron" is an intervening sequence in a gene that does not encode a portion of the protein sequence. Thus, such sequences are transcribed into RNA but are then excised and are not translated. The term is also used for the excised RNA sequences.

An "exon" is a portion of the sequence of a gene that is transcribed and is found in the mature messenger RNA derived from the gene, but is not necessarily a part of the sequence that encodes the final gene product.

The term "deduced nucleotide sequence" refers to a DNA sequence after removal of intervening sequences, based on homology to other DNA sequences encoding the same protein.

The term "deduced amino acid sequence" refers to a polypeptide sequence derived from a DNA sequence after removal of intervening sequences, based on homology to other proteins encoded by DNA sequences encoding the same protein.

The term "translation leader sequence" refers to a DNA sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D. (1995) *Molecular Biotechnology* 3:225).

The "3' non-coding sequences" refer to DNA sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al., (1989) *Plant Cell* 1:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from post-transcriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by

the cell. "cDNA" refers to a DNA that is complementary to and synthesized from a mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into the double-stranded form using the Klenow fragment of DNA polymerase I. "Sense" RNA refers to RNA transcript that includes the mRNA and can be translated into protein within a cell or *in vitro*. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target isolated nucleic acid fragment (U.S. Patent No. 5,107,065). The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes. The terms "complement" and "reverse complement" are used interchangeably herein with respect to mRNA transcripts, and are meant to define the antisense RNA of the message.

The term "endogenous RNA" refers to any RNA which is encoded by any nucleic acid sequence present in the genome of the host, whether naturally-occurring or non-naturally occurring, i.e., introduced by recombinant means, mutagenesis, etc.

The term "non-naturally occurring" means artificial, not consistent with what is normally found in nature.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a coding sequence when it is capable of regulating the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in a sense or antisense orientation. In another example, the complementary RNA regions of the invention can be operably linked, either directly or indirectly, 5' to the target mRNA, or 3' to the target mRNA, or within the target mRNA, or a first complementary region is 5' and its complement is 3' to the target mRNA.

The term "expression", as used herein, refers to the production of a functional end-product. Expression of an isolated nucleic acid fragment involves transcription of the isolated nucleic acid fragment and translation of the mRNA into a precursor or mature protein. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020).

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

"Stable transformation" refers to the transfer of a nucleic acid fragment into a genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance. In contrast, "transient transformation" refers to the transfer of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without integration or stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. The preferred method of cell transformation of rice, corn and other monocots is the use of particle-accelerated or "gene gun" transformation technology (Klein et al., (1987) *Nature (London)* 327:70-73; U.S. Patent No. 4,945,050), or an *Agrobacterium*-mediated method using an appropriate Ti plasmid containing the transgene (Ishida Y. et al., 1996, *Nature Biotech.* 14:745-750). The term "transformation" as used herein refers to both stable transformation and transient transformation.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Sambrook").

The term "recombinant" refers to an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques.

"PCR" or "Polymerase Chain Reaction" is a technique for the synthesis of large quantities of specific DNA segments, consists of a series of repetitive cycles (Perkin Elmer Cetus Instruments, Norwalk, CT). Typically, the double stranded DNA is heat denatured, the two primers complementary to the 3' boundaries of the target segment are annealed at low temperature and then extended at an intermediate temperature. One set of these three consecutive steps is referred to as a cycle.

Polymerase chain reaction ("PCR") is a powerful technique used to amplify DNA millions of fold, by repeated replication of a template, in a short period of time. (Mullis et al, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273 (1986); Erlich et al, European Patent Application 50,424; European Patent Application 84,796; European Patent Application 258,017, European Patent Application 237,362; Mullis,

European Patent Application 201,184, Mullis et al U.S. Patent No. 4,683,202; Erlich, U.S. Patent No. 4,582,788; and Saiki et al, U.S. Patent No. 4,683,194). The process utilizes sets of specific in vitro synthesized oligonucleotides to prime DNA synthesis. The design of the primers is dependent upon the sequences of DNA that are desired to be analyzed. The technique is carried out through many cycles (usually 20-50) of melting the template at high temperature, allowing the primers to anneal to complementary sequences within the template and then replicating the template with DNA polymerase.

The products of PCR reactions are analyzed by separation in agarose gels followed by ethidium bromide staining and visualization with UV transillumination. Alternatively, radioactive dNTPs can be added to the PCR in order to incorporate label into the products. In this case the products of PCR are visualized by exposure of the gel to x-ray film. The added advantage of radiolabeling PCR products is that the levels of individual amplification products can be quantitated.

The terms "recombinant construct", "expression construct" and "recombinant expression construct" are used interchangeably herein. These terms refer to a functional unit of genetic material that can be inserted into the genome of a cell using standard methodology well known to one skilled in the art. Such construct may be itself or may be used in conjunction with a vector. If a vector is used then the choice of vector is dependent upon the method that will be used to transform host plants as is well known to those skilled in the art. For example, a plasmid vector can be used. The skilled artisan is well aware of the genetic elements that must be present on the vector in order to successfully transform, select and propagate host cells comprising any of the isolated nucleic acid fragments of the invention. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al., (1985) *EMBO J.* 4:2411-2418; De Almeida et al., (1989) *Mol. Gen. Genetics* 218:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

Co-suppression constructs in plants previously have been designed by focusing on overexpression of a nucleic acid sequence having homology to an endogenous mRNA, in the sense orientation, which results in the reduction of all RNA having homology to the overexpressed sequence (see Vaucheret et al. (1998) *Plant J* 16:651-659; and Gura (2000) *Nature* 404:804-808). The overall efficiency of this phenomenon is low, and the extent of the RNA reduction is widely variable. Recent work has described the use of "hairpin" structures that incorporate all, or

part, of an mRNA encoding sequence in a complementary orientation that results in a potential "stem-loop" structure for the expressed RNA (PCT Publication WO 99/53050 published on October 21, 1999). This increases the frequency of co-suppression in the recovered transgenic plants. Another variation describes the use of plant viral sequences to direct the suppression, or "silencing", of proximal mRNA encoding sequences (PCT Publication WO 98/36083 published on August 20, 1998). Both of these co-suppressing phenomena have not been elucidated mechanistically, although recent genetic evidence has begun to unravel this complex situation (Elmayan et al. (1998) *Plant Cell* 10:1747-1757).

In one aspect, this invention includes an isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide required for proper root formation, wherein the polypeptide has an amino acid sequence of at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity, based on the Clustal V method of alignment, when compared to one of SEQ ID NO:6, 8, 30, or 38. The polypeptide may also comprise SEQ ID NO:6, 8, 30, or 38, and the nucleotide sequence may comprise SEQ ID NO:5, 7, 29, or 37.

Also included in the present invention is a complement of any of the foregoing nucleotide sequences, wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

As used herein, "proper root formation" means a root system that exhibits formation of all nodal roots (lateral seminal, crown, and brace roots). A polypeptide is required for proper root formation in that the absence of or altered levels of the polypeptide in a plant results in the elimination of or altered levels of formation of one or more of the nodal roots.

In another aspect, this invention includes isolated polynucleotides as described herein (or complements), wherein the nucleotide sequence comprises at least two, three, four, or five motifs selected from group consisting of SEQ ID NOs:9, 10, 11, 12 and 13, wherein said motif is a substantially conserved subsequence.

"Motifs" or "subsequences" refer to short regions of conserved sequences of nucleic acids or amino acids that comprise part of a longer sequence. For example, it is expected that such conserved subsequences (for example SEQ ID NOs: 9,10,11,12 and 13) would be important for function, and could be used to identify new homologues of RTCS-homologues in plants. It is expected that some or all of the elements may be found in an RTCS-homologue. Also, it is expected that at least one or two of the conserved amino acids in any given motif may differ in a true RTCS-homologue.

In another aspect, a polynucleotide of this invention or a functionally equivalent subfragment thereof is useful in antisense inhibition or cosuppression of

expression of nucleic acid sequences encoding proteins required for proper root formation, most preferably in antisense inhibition or cosuppression of an endogenous RTCS or heterologous RTCS gene.

5 Protocols for antisense inhibition or co-suppression are well known to those skilled in the art and are described above.

In still a further aspect, this invention includes an isolated nucleic acid fragment comprising (a) a promoter consisting essentially of SEQ ID NO:2, 3 or 4, or (b) a substantially similar and functionally equivalent subfragment of said promoter.

10 Also of interest are recombinant DNA constructs comprising any of the above-identified isolated nucleic acid fragments or isolated polynucleotides or complements thereof or parts of such fragments or complements, operably linked to at least one regulatory sequence.

Plants, plant tissue or plant cells comprising such recombinant DNA constructs in their genome are also within the scope of this invention.

15 Transformation methods are well known to those skilled in the art and are described above. Any plant, dicot or monocot can be transformed with such recombinant DNA constructs.

20 Examples of monocots include, but are not limited to, corn, wheat, rice, sorghum, millet, barley, palm, lily, *Alstroemeria*, rye, and oat. Examples of dicots include, but are not limited to, soybean, rape, sunflower, canola, grape, guayule, columbine, cotton, tobacco, peas, beans, flax, safflower, alfalfa.

25 Plant tissue includes differentiated and undifferentiated tissues or plants, including but not limited to, roots, stems, shoots, leaves, pollen, seeds, tumor tissue, and various forms of cells and culture such as single cells, protoplasm, embryos, and callus tissue. The plant tissue may in plant or in organ, tissue or cell culture.

In another aspect, this invention includes a method of altering root structure during plant development, comprising:

- 30 (a) transforming a plant with a recombinant DNA construct of the invention;  
(b) growing the transformed plant under conditions suitable for the expression of the recombinant DNA construct; and  
(c) selecting those transformed plants having altered root structure.

35 As used herein, altering root structure includes altering one or more of the nodal roots (lateral seminal, crown, and brace roots). Alterations may include alterations in the level of growth of any or all of the nodal roots or alterations in root architecture. The alterations may result in increased or decreased changes.

The regeneration, development, and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: Methods for Plant Molecular Biology, (Eds.),

Academic Press, Inc. San Diego, CA, (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated.

- 5 The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous isolated nucleic acid fragment that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous  
10 transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

- 15 There are a variety of methods for the regeneration of plants from plant tissue.

The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

- Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens*, and obtaining transgenic plants have been published for cotton (U.S. Patent No. 5,004,863, U.S. Patent No. 5,159,135, U.S. Patent No. 5,518, 908); soybean (U.S. Patent No. 5,569,834, U.S. Patent No. 5,416,011, McCabe et. al., *BiolTechnology* 6:923 (1988), Christou et al., *Plant Physiol.* 87:671-674 (1988)); *Brassica* (U.S. Patent No. 5,463,174); peanut (Cheng et al., *Plant Cell Rep.* 15:653-657 (1996), McKently et al., *Plant Cell Rep.* 14:699-703 (1995)); papaya; and pea (Grant et al., *Plant Cell Rep.* 15:254-258, (1995)).

- Transformation of monocotyledons using electroporation, particle bombardment, and *Agrobacterium* have also been reported. Transformation and plant regeneration have been achieved in asparagus (Bytebier et al., *Proc. Natl. Acad. Sci. (USA)* 84:5354, (1987)); barley (Wan and Lemaux, *Plant Physiol* 104:37 (1994)); *Zea mays* (Rhodes et al., *Science* 240:204 (1988), Gordon-Kamm et al., *Plant Cell* 2:603-618 (1990), Fromm et al., *BiolTechnology* 8:833 (1990), Koziel et al., *BiolTechnology* 11: 194, (1993), Armstrong et al., *Crop Science* 35:550-557 (1995)); oat (Somers et al., *BiolTechnology* 10: 15 89 (1992)); orchard grass (Horn et al., *Plant Cell Rep.* 7:469 (1988)); rice (Toriyama et al., *TheorAppl. Genet.* 205:34, (1986); Part et al., *Plant Mol. Biol.* 32:1135-1148, (1996); Abedinia et al., *Aust. J. Plant Physiol.* 24:133-141 (1997); Zhang and Wu, *Theor. Appl. Genet.* 76:835 (1988); Zhang et al. *Plant Cell Rep.* 7:379, (1988); Battraw and Hall, *Plant*



*Sci.* 86:191-202 (1992); Christou et al., *Bio/Technology* 9:957 (1991)); rye (De la Pena et al., *Nature* 325:274 (1987)); sugarcane (Bower and Birch, *Plant J.* 2:409 (1992)); tall fescue (Wang et al., *Bio/Technology* 10:691 (1992)), and wheat (Vasil et al., *Bio/Technology* 10:667 (1992); U.S. Patent No. 5,631,152).

5        Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte et al., *Nature* 335:454-457 (1988); Marcotte et al., *Plant Cell* 1:523-532 (1989); McCarty et al., *Cell* 66:895-905 (1991); Hattori et al.,  
10    *Genes Dev.* 6:609-618 (1992); Goff et al., *EMBO J.* 9:2517-2522 (1990)).

      Transient expression systems may be used to functionally dissect isolated nucleic acid fragment constructs (see generally, Maliga et al., *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995)). It is understood that any of the nucleic acid molecules of the present invention can be introduced into a plant  
15    cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers etc.

      In addition to the above discussed procedures, practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of macromolecules (e.g., DNA  
20    molecules, plasmids, etc.), generation of recombinant organisms and the screening and isolating of clones, (see for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989); Maliga et al., *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995); Birren et al., *Genome Analysis: Detecting Genes*, 1, Cold Spring Harbor, New York (1998); Birren et al.,  
25    *Genome Analysis: Analyzing DNA*, 2, Cold Spring Harbor, New York (1998); *Plant Molecular Biology: A Laboratory Manual*, eds. Clark, Springer, New York (1997)).

      In a still further aspect, this invention includes a method to isolate nucleic acid fragments encoding polypeptides associated with altering root structure during plant development, which comprises:

- 30        (a) comparing SEQ ID NOs: 6, 8, 30, or 38 with other polypeptide sequences associated with altering root structure during plant development;
- (b) identifying the conserved sequences(s) of 4 or more amino acids obtained in step (a);
- (c) making region-specific nucleotide probe(s) or oligomer(s) based on the  
35    conserved sequences identified in step (b); and
- (d) using the nucleotide probe(s) or oligomer(s) of step (c) to isolate sequences associated with altering root structure during plant development by sequence dependent protocols.

Examples of conserved sequence elements that would be useful in identifying other plant sequences associated with altering root structure during plant development can be found in the group comprising, but not limited to, the nucleotides encoding the polypeptides of SEQ ID NO:9, 10,11, 12, and 13.

5 In another aspect, this invention also includes a method of mapping genetic variations related to altering root structure during plant development comprising:

- (a) crossing two plant varieties; and
- (b) evaluating genetic variations with respect to:
  - 10 (i) a nucleic acid sequence selected from the group consisting of SEQ ID NO:5,7, 29, or 37; or
  - (ii) a nucleic acid sequence encoding a polypeptide selected from the group consisting of SEQ ID NO: 6, 8 ,30, or 38 in progeny plants resulting from the cross of step (a) wherein the evaluation is made using a method selected from the group consisting of: RFLP
  - 15 analysis, SNP analysis, and PCR-based analysis.

In another embodiment, this invention includes a method of molecular breeding to obtain altered root formation:

- (a) crossing two plant varieties; and
- (b) evaluating genetic variations with respect to:
  - 20 (i) a nucleic acid sequence selected from the group consisting of SEQ ID NO: 5,7 ,29, or 37; or
  - (ii) a nucleic acid sequence encoding a polypeptide selected from the group consisting of SEQ ID NO: 6,8, 30, or 38
  - in progeny plants resulting from the cross of step (a) wherein the
  - 25 evaluation is made using a method selected from the group consisting of: RFLP analysis, SNP analysis, and PCR-based analysis.

The terms "mapping genetic variation" or "mapping genetic variability" are used interchangeably and define the process of identifying changes in DNA sequence, whether from natural or induced causes, within a genetic region that

30 differentiates between different plant lines, cultivars, varieties, families, or species. The genetic variability at a particular locus (gene) due to even minor base changes can alter the pattern of restriction enzyme digestion fragments that can be generated. Pathogenic alterations to the genotype can be due to deletions or

35 insertions within the gene being analyzed or even single nucleotide substitutions that can create or delete a restriction enzyme recognition site. RFLP analysis takes advantage of this and utilizes Southern blotting with a probe corresponding to the isolated nucleic acid fragment of interest.

Thus, if a polymorphism (i.e., a commonly occurring variation in a gene or segment of DNA; also, the existence of several forms of a gene (alleles) in the same species) creates or destroys a restriction endonuclease cleavage site, or if it results in the loss or insertion of DNA (e.g., a variable nucleotide tandem repeat (VNTR) polymorphism), it will alter the size or profile of the DNA fragments that are generated by digestion with that restriction endonuclease. As such, individuals that possess a variant sequence can be distinguished from those having the original sequence by restriction fragment analysis. Polymorphisms that can be identified in this manner are termed "restriction fragment length polymorphisms: ("RFLPs").

RFLPs have been widely used in human and plant genetic analyses (Glassberg, UK Patent Application 2135774; Skolnick et al, *Cytogen. Cell Genet.* 32:58-67 (1982); Botstein et al, *Ann. J. Hum. Genet.* 32:314-331 (1980); Fischer et al (PCT Application WO 90/13668; Uhlen, PCT Application WO 90/11369).

A central attribute of "single nucleotide polymorphisms" or "SNPs" is that the site of the polymorphism is at a single nucleotide. SNPs have certain reported advantages over RFLPs or VNTRs. First, SNPs are more stable than other classes of polymorphisms. Their spontaneous mutation rate is approximately  $10^{-9}$  (Kornberg, *DNA Replication*, W.H. Freeman & Co., San Francisco, 1980), approximately, 1,000 times less frequent than VNTRs (U.S. Patent 5,679,524).

Second, SNPs occur at greater frequency, and with greater uniformity than RFLPs and VNTRs. As SNPs result from sequence variation, new polymorphisms can be identified by sequencing random genomic or cDNA molecules. SNPs can also result from deletions, point mutations and insertions. Any single base alteration, whatever the cause, can be a SNP. The greater frequency of SNPs means that they can be more readily identified than the other classes of polymorphisms.

SNPs can be characterized using any of a variety of methods. Such methods include the direct or indirect sequencing of the site, the use of restriction enzymes where the respective alleles of the site create or destroy a restriction site, the use of allele-specific hybridization probes, the use of antibodies that are specific for the proteins encoded by the different alleles of the polymorphism or by other biochemical interpretation. SNPs can be sequenced by a number of methods. Two basic methods may be used for DNA sequencing, the chain termination method of Sanger et al, *Proc. Natl. Acad. Sci. (U.S.A.)* 74:5463-5467 (1977), and the chemical degradation method of Maxam and Gilbert, *Proc. Natl. Acad. Sci. (U.S.A.)* 74: 560-564 (1977).

Furthermore, single point mutations can be detected by modified PCR techniques such as the ligase chain reaction ("LCR") and PCR-single strand conformational polymorphisms ("PCR-SSCP") analysis. The PCR technique can

also be used to identify the level of expression of genes in extremely small samples of material, e.g., tissues or cells from a body. The technique is termed reverse transcription-PCR ("RT-PCR").

5 The term "molecular breeding" defines the process of tracking molecular markers during the breeding process. It is common for the molecular markers to be linked to phenotypic traits that are desirable. By following the segregation of the molecular marker or genetic trait, instead of scoring for a phenotype, the breeding process can be accelerated by growing fewer plants and eliminating assaying or visual inspection for phenotypic variation. The molecular markers useful in this  
10 process include, but are not limited to, any marker useful in identifying mapable genetic variations previously mentioned, as well as any closely linked genes that display synteny across plant species. The term "synteny" refers to the conservation of gene placement/order on chromosomes between different organisms. This means that two or more genetic loci, that may or may not be closely linked, are  
15 found on the same chromosome among different species. Another term for synteny is "genome colinearity".

### EXAMPLES

20 The present invention is further defined in the following Examples, in which parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope  
25 thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

30

### EXAMPLE 1

#### Map-based Cloning of RTCS

35 In order to map the *rtcs* mutation, two mapping populations and their corresponding corn seeds, segregating for the *rtcs* gene, were utilized. The first mapping populations consisted of 1591 plants derived from a F1 cross between inbred line DK105, carrying the *rtcs* mutation, and the inbred line B73. The second mapping populations consisted of 475 plants derived from a F1 cross between inbred line DK105, carrying the *rtcs* mutation, and the inbred line Mo17.

Homozygous *rtcs/rtcs* plants were scored as completely lodged plants when grown in the field for 40 days or more. In the B73 and Mo17 F2 populations, 388 and 61 mutant plants were retrieved, respectively. These plants were selected for fine mapping of the *rtcs* locus.

5 DNA was extracted from those plants using standard molecular biology procedures.

To obtain F2 plants that carry recombination near the *rtcs* locus, public PCR-based DNA markers (SSRs) present in the Maize Genetics and Genomic Database (MaizeGDB), were used. When these were not available, CAP (allele-specific PCR  
10 primers) markers were developed from the DuPont proprietary sequences of BAC (Bacterial Artificial Chromosome) clones of known map positions. Both CAP and SSR primers were used in a PCR reaction containing 25ng of DNA.

15 Flanking SSR marker BNLG 1014 [BNLG 1014 forward primer (SEQ ID NO:14), BNLG 1014 reverse primer (SEQ ID NO: 15)] and BNLG 1429 [BNLG 1429 forward primer (SEQ ID NO:16), BNLG 1429 reverse primer (SEQ ID NO: 17)] were retrieved from the MaizeGDB. These markers are localized at 82.80 cM and 143.50 cM of Chromosome 1 respectively, based on the public map IBM2 neighbors 1.

20 SSR markers were analyzed using a Perkin Elmer ABI 3700 machine. PCR was performed in 20 ul, using Qiagen Hot start mix (Qiagen), following the manufacturer instructions, and using one of the two amplifying primers labeled with a specific fluorochrome.

When using these 2 primers on 367 *rtcs* plants, a total of 97 recombinants were obtained, 34 with marker BNLG 1014 and 63 from marker BNLG1429, indicating that *rtcs* was closer to BNLG 1014.

25 In order to obtain genetic markers closer to *rtcs*, more primers were retrieved from the Maize GDB based on their position along chromosome 1 and tested on the same number of individuals. In particular, markers UMC1685 [UMC1685 forward primer (SEQ ID NO: 18), UMC 1685 reverse primer (SEQ ID NO: 19)] gave 7 recombinants and marker UMC1660 [UMC1660 forward primer (SEQ ID NO:20),  
30 UMC 1660 reverse primer (SEQ ID NO:21)] gave 11 recombinants indicating a distance of 0.95 cM and 1.5 cM from the *rtcs* locus respectively.

35 Marker UMC1685 and UMC 1660 have been physically positioned by hybridization onto a single maize contig, named 1871 (Dupont Genomix database). The physical distance between the two markers encompasses approximately 10 BACs.

Based on this information, new CAP markers were designed using available BAC-end sequences of the BACs constituting the region of contig 1871 surrounded by markers UMC 1685 and 1660.

Cap marker b104.i24 [b104.i24 forward primer (SEQ ID NO:22), b104.i24 reverse primer (SEQ ID NO:23)] was designed based on the BAC-end sequence of clone BAC 104h.i24. This primer set amplifies a region of 450 bp, showing polymorphism between B73 and DK105 after restriction with the 4-cutter enzyme HhaI.

CAP marker amplifications were performed in a 25 ul PCR reaction using the Qiagen HotStart mix and 25 ng DNA. The thermal cycle conditions were: 95°C 15min (1 cycle), 94°C 45 sec, 56°C 45 sec, 72°C 45 sec, (35 cycles) 72°C 7 min.

3 ul of the amplification product was used for a restriction digest (total volume of 15 ul) with the 4-cutter restriction enzyme HhaI (Promega). Restriction reaction was carried out at 37°C for one hour. Restricted amplification products were examined on 2.5 % agarose gels. By screening the 18 previously obtained recombinants with this primers set, only 7 recombination breakpoint were found, on the same side of the marker UMC1160, meaning that *rtcs* lies exactly in the middle between markers UMC 1685 and b104.124.

Cap marker b74.m9 [ b74.m9 forward primer (SEQ ID NO:24), b74.m9 reverse primer (SEQ ID NO:25)] was designed based on the BAC-end sequence of clone BAC b74a.m9. This primer set amplifies a region of 313 bp, showing polymorphism between B73 and DK105 after restriction with the 4-cutter enzyme HaeIII. This CAP marker allowed us to narrow down the genomic region containing the *rtcs* locus between marker UMC 1685 (at a distance of 0.95 cM, with 7 recombinant plants), located on BAC clone 35.m15 and marker b74.m9, corresponding to the end of the BAC clone b74.m9 at a distance of 0.13 cM (one recombination breakpoint). The 2 BAC clones are overlapping.

## EXAMPLE 2

### Identification of the *RTCS* Gene

In order to identify the *RTCS* gene that was mapped to the region comprising the two overlapping BAC clones, BAC 74.m9 was sequenced. The 6kb fragment of Bac74.m9 containing the *RTCS* gene is shown in SEQ ID NO.:1. For the purpose, BAC DNA was nebulized using high-pressure nitrogen gas as described in Roe et al. 1996 (Roe et al. (1996) "DNA isolation and Sequencing" John Wiley and Sons, New York).

The estimated 150 Kb of sequence of BAC 74.m9 was searched for the presence of open reading frames, and 4 regions, showing similarities to genes filed in Genbank as well as in the DuPont proprietary EST database, were identified. These regions correspond to the rice genes annotated as OSJNBb0050N02.6, OSJNBb0050N02.7, OSJNBb0050N02.9 and OSJNBb0050N02.10 of the rice BAC

OSJNBb0050N02, annotated in Genbank as AC105734. These rice genes have been annotated as "hypothetical proteins". The corresponding corn ESTs are cds3f.pk004.j15, cen3n.pk0159.d12, cr1n.pk0028.h3a:fis and cpf1c.pk006.d18a:fis (from DuPont proprietary EST database). These genes were selected because they were in the middle of the recombination data.

The candidate genes were then PCR amplified from genomic DNA of the wild type (DK105) and the mutant genotypes (*rtcs*), and the sequences compared. By comparing the sequences of (1) a gene encoding a corn homologue (amplified by PCR from the wild type and the mutant using the sequence from bac 74.m9 (SEQ ID NO:1) corresponding to the rice gene OSJNBb0050N02.10 (which is annotated as "hypothetical protein" in Genbank) and (2) one of the arabidopsis "LOB domain" gene family members (GI No. 2761471) a mutation was found in the mutant allele. The *rtcs* allele carries a 5bp insertion at position 227bp from the starting ATG, which causes a frameshift and introduces a premature stop codon.

### EXAMPLE 3

#### Cloning the RTCS cDNA

Total RNA can be extracted from developing maize using a TRIzol® Reagent obtained from Life Technologies Inc., Rockville, MD, 20849 (GIBCO-BRL) that contains phenol and guanidine thiocyanate. Poly A mRNA can be purified from total RNA with mRNA Purification kits obtained from Amersham Pharmacia Biotech Inc., Piscataway, NJ, 08855, which consists of oligo (dT)-cellulose spin columns. To make the cDNA library, 5.5 ug of polyA RNA can be used for cDNA synthesis kits, which can be obtained from Stratagene, La Jolla, CA, 92037. Superscript® reverse transcriptase can be obtained from Life Technologies Inc., Rockville, MD, 20849 (GIBCO-BRL). BRL cDNA Size Fraction Columns (GIBCO-BRL) can be used to fractionate the cDNA by size, fractions can be precipitated, resuspended and ligated with 1 ug of the Uni-ZAP XR vector. After ligation it can be packaged in Gigapack III Gold® packaging extract obtained from Stratagene, La Jolla, CA, 92037. The unamplified library titer can be estimated. An appropriate amount can be used for amplification purposes to produce amplified cDNA.

Screening for the *RTCS* cDNA follows standard protocols well known to those skilled in the art (Ausubel et al. 1993, "Current Protocols in Molecular Biology" John Wiley & Sons, USA, or Sambrook et al. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press). Briefly,  $1.5 \times 10^6$  phage

clones can be plated, then transferred to nylon membranes, which then will be subjected to hybridization with radioactively labeled *RTCS* probe. Positives are isolated and examined for their identity as *RTCS* cDNAs through PCR with *RTCS*-specific primers. The longest cDNA clones that give positive results from the PCR reaction are isolated and sequenced.

#### EXAMPLE 4

##### Cloning and isolation of a full length *RTCS* cDNA clone from corn

A lambda cDNA library (cdr1f) was prepared from 15 days old B73 nodal roots. The library was screened using a radioactive probe generated by PCR amplification of the bac clone BAC74.m9 using the the primers shown in SEQ ID NO.: 31 and 32 and as described in Example 3. The full length *RTCS* cDNA clone was retrieved after screening  $1.8 \times 10^6$  lambda clones. The sequence of the full length corn *RTCS* cDNA clone is identical to SEQ ID NO:29, except for one nucleotide difference which does not change the encoded amino acid sequence, which is identical to SEQ ID NO:30.

#### EXAMPLE 5

##### Genetic Confirmation of the *RTCS* gene

The genetic confirmation that the *RTCS* isolated nucleic acid fragment encodes the polypeptide responsible for altering root structure can be accomplished by transforming *rtcs* mutants with the isolated *RTCS* cloned sequence.

*RTCS* homologs from other crop species can also be tested in this system by obtaining full-gene sequences, ligation to an appropriate promoter, such as the *RTCS* promoter and complementing the maize *RTCS* mutant.

In order to confirm possible tissue-specific expression of the *RTCS* gene, the presence of the *RTCS* transcript in various tissues can be analyzed by RNA blot analysis and in situ hybridization.

One method for transforming DNA into cells of higher plants that is available to those skilled in the art is high-velocity ballistic bombardment using metal particles coated with the nucleic acid constructs of interest (see Klein et al. Nature (1987) (London) 327:70-73, and see U.S. Patent No. 4,945,050). A Biolistic PDS-1000/He (BioRAD Laboratories, Hercules, CA) can be used for these complementation experiments (see Example 4 for further details). The particle bombardment technique can be used to transform the *rtcs* mutant with the cloned *RTCS* wild type sequence [SEQ ID NO:5 or 7], encoding a functional *RTCS* protein.

The bacterial hygromycin B phosphotransferase (Hpt II) gene from *Streptomyces hygrosopicus* that confers resistance to the antibiotic hygromycin



can be used as the selectable marker for the maize transformation. In the vector, pML18, the Hpt II gene can be engineered with the 35S promoter from Cauliflower Mosaic Virus and the termination and polyadenylation signals from the octopine synthase gene of *Agrobacterium tumefaciens*. pML18 was described in  
5 WO 97/47731, which was published on December 18, 1997, the disclosure of which is hereby incorporated by reference.

Embryogenic maize callus cultures derived serve as source material for transformation experiments. This material can be generated by germinating sterile maize seeds on a callus initiation media (MS salts, Nitsch and Nitsch vitamins,  
10 1.0 mg/l 2,4-D and 10  $\mu$ M AgNO<sub>3</sub>) in the dark at 27-28°C. Embryogenic callus proliferating from the scutellum of the embryos is then transferred to CM media (N6 salts, Nitsch and Nitsch vitamins, 1 mg/l 2,4-D, Chu et al., 1985, *Sci. Sinica* 18: 659-668). Callus cultures are maintained on CM by routine sub-culture at two week intervals and used for transformation within 10 weeks of initiation.

15 Callus can be prepared for transformation by subculturing 0.5-1.0 mm pieces approximately 1 mm apart, arranged in a circular area of about 4 cm in diameter, in the center of a circle of Whatman #541 paper placed on CM media. The plates with callus are incubated in the dark at 27-28°C for 3-5 days. Prior to bombardment, the filters with callus are transferred to CM supplemented with 0.25 M mannitol and  
20 0.25 M sorbitol for 3 hr in the dark. The petri dish lids are then left ajar for 20-45 minutes in a sterile hood to allow moisture on tissue to dissipate.

Each genomic DNA fragment is co-precipitated with pML18 containing the selectable marker for maize transformation onto the surface of gold particles. To accomplish this, a total of 10  $\mu$ g of DNA at a 2:1 ratio of trait:selectable marker  
25 DNAs are added to 50  $\mu$ l aliquot of gold particles that are resuspended at a concentration of 60 mg ml<sup>-1</sup>. Calcium chloride (50  $\mu$ l of a 2.5 M solution) and spermidine (20  $\mu$ l of a 0.1 M solution) are then added to the gold-DNA suspension as the tube was vortexed for 3 min. The gold particles are centrifuged in a microfuge for 1 sec and the supernatant removed. The gold particles are then  
30 washed twice with 1 ml of absolute ethanol and then resuspended in 50 ml of absolute ethanol and sonicated (bath sonicator) for one second to disperse the gold particles. The gold suspension is incubated at -70°C for five minutes and sonicated (bath sonicator) if needed to disperse the particles. Six  $\mu$ l of the DNA-coated gold particles are then loaded onto mylar macrocarrier disks and the ethanol is allowed to  
35 evaporate.

At the end of the drying period, a petri dish containing the tissue is placed in the chamber of the PDS-1000/He. The air in the chamber is then evacuated to a vacuum of 28-29 inches Hg. The macrocarrier is accelerated with a helium shock

wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1080-1100 psi. The tissue is placed approximately 8 cm from the stopping screen and the callus was bombarded two times. Two to four plates of tissue are bombarded in this way with the DNA-coated gold particles. Following

5 bombardment, the callus tissue is transferred to CM media without supplemental sorbitol or mannitol.

Within 3-5 days after bombardment the callus tissue is transferred to SM media (CM medium containing 50 mg/l hygromycin). To accomplish this, callus tissue is transferred from plates to sterile 50 ml conical tubes and weighed. Molten  
10 top-agar at 40° C is added using 2.5 ml of top agar/100 mg of callus. Callus clumps are broken into fragments of less than 2 mm diameter by repeated dispensing through a 10 ml pipet. Three ml aliquots of the callus suspension are plated onto fresh SM media and the plates are incubated in the dark for 4 weeks at 27-28°C. After 4 weeks, transgenic callus events are identified, transferred to fresh SM plates  
15 and grown for an additional 2 weeks in the dark at 27-28°C.

Growing callus can then be transferred to RM1 media (MS salts, Nitsch and Nitsch vitamins, 2% sucrose, 3% sorbitol, 0.4% gelrite +50 ppm hyg B) for 2 weeks in the dark at 25°C. After 2 weeks the callus can be transferred to RM2 media (MS salts, Nitsch and Nitsch vitamins, 3% sucrose, 0.4% gelrite + 50 ppm hyg B) and  
20 placed under cool white light ( $\sim 40 \mu\text{Em}^{-2}\text{s}^{-1}$ ) with a 12 hr photoperiod at 25°C and 30-40% humidity. After 2-4 weeks in the light, callus can begin to organize, and form shoots. Shoots can be removed from surrounding callus/media and gently transferred to RM3 media (1/2 x MS salts, Nitsch and Nitsch vitamins, 1% sucrose + 50 ppm hygromycin B) in phytatrays (Sigma Chemical Co., St. Louis, MO) and  
25 incubation can be continued using the same conditions as described in the previous step.

Plants can then be transferred from RM3 to 4" pots containing Metro mix 350 after 2-3 weeks, when sufficient root and shoot growth has occurred. The seed obtained from the transgenic plants can be examined for genetic complementation  
30 of the *rtcs* mutation with the wild-type genomic DNA containing the *RTCS* gene.

### EXAMPLE 6

#### Complementation with the corn RTCS and rice RTCS genes

Two constructs are made and used for transformation into corn callus derived  
35 from a cross between the inbred line GS3 and the mutant *rtcs* line, in order to confirm the function of the corn RTCS and the rice RTCS gene.

Construct 1: BAC 74.m9 is digested with the restriction enzymes XmaI and HindIII. After gel separation on a 0.7% low melting agarose gel, a fragment of 5408

bp containing the corn RTCS gene, including 2410 bp of endogeneous promoter and 2166 bp of sequence at the 3' end of the RTCS ORF, is recovered and used for ligation into vector PHP20067 digested with XmaI and HindIII, resulting in vector PHP24451 (Fig. 2). The 5408 bp sequence extends from nucleotides 369 to 5776 of SEQ ID NO: 1.

Construct 2: The rice RTCS ORF (SEQ ID NO: 7 ) is used to substitute the corn ORF in the 5408 bp fragment, but the corn promoter region and 3' sequence are maintained in the construct. The construct is prepared as follows. Rice genomic DNA is amplified using the primers shown in SEQ ID NO.: 33 and 34. The PCR fragment is cut with the restriction enzymes XhoI and DralI and purified. The corn RTCS 5408bpXmaI- HindIII fragment is cut with XhoI and DralI, and the 3 pieces (corn XmaI- XhoI, rice XhoI-DraI, corn DraI-HindIII) are ligated in a equimolar reaction. The resulting chimera is cloned into vector PHP20067 digested with XmaI and HindIII, resulting in vector PHP 24452 (Fig. 3) .

Vector PHP20067 is constructed using standard molecular biology techniques used by those of skill in the art (Sambrook et al., supra). The plasmid pSB11 can be obtained from Japan Tobacco Inc. (Tokyo, Japan). The construction of pSB11 from pSB21 and the construction of pSB21 from starting vectors is described by Komari et al. (1996, Plant J. 10:165-174). The plasmid pSB1 (Japan Tobacco Inc., Tokyo, Japan) can further be modified by the addition of a multiple cloning site and a selectable marker gene directing herbicide resistance in plants and plant tissues. In the case of PHP20067, this selectable marker gene comprises a promoter/Intron from the UBIQUITIN gene of *Z. mays*, a coding sequence for MO-PAT (a maize optimized version of the phosphinothricin acetyltransferase gene from *Streptomyces viridochromogenes*) and a polyadenylation/terminator sequence (PINII TERM) from potato. This gene cassette is ligated into the pSB11-derived vector between the newly introduced multiple cloning site and the LB (Left Border) to generate PHP20067.

After selfing, the regenerated plants will segregate for the presence of the transgene. Only rtcs mutant plants expressing the corn RTCS or rice RTCS gene will show a root phenotype identical to wild type.

### EXAMPLE 7

#### Identification of Genomic and cDNA Clones

A maize inbred B73 genomic DNA assembly for the putative RTCS gene was created using BLAST search (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410) with the maize genomic RTCS sequence from

Mo17, contained in SEQ ID NO:1, conducted against the GSS (Genomic Survey Sequences) datasets available in GenBank.

Selected matches (approximately 19 GSS fragments) were downloaded as HTML/text files for import into a sequence assembly program (Sequencher v4.1.4b).

5 Once trimmed of HTML code, these individual fragments were assembled using standard parameters in Sequencher. The contig alignment was then examined to make edits to the consensus sequence. Several of the 19 imported fragments did not assemble and were rejected.

10 The 13 remaining fragments resulted in a RTCS contig of 3286bp which is shown in SEQ ID NO:28. The accession numbers of the 13 fragments are :  
gi:32081270, gi:31973186, gi:34279177, gi:34277557, gi:34279170, gi:34277545,  
gi:32081279, gi:31973192, gi:33915405, gi:33915408, gi:34051270, gi:34051273,  
gi:34051273.

15 Clones for cDNAs encoding RTCS-like proteins are identified by conducting BLAST searches. (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ  
20 databases). The sequence obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database  
25 using the BLASTX algorithm (Gish and States (1993) *Nat. Genet.* 3:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly,  
30 the greater the pLog value, the greater the likelihood that the sequence and the BLAST "hit" represent homologous proteins.

ESTs submitted for analysis are compared to the genbank database as described above. ESTs that contain sequences more 5- or 3-prime can be found by using the BLASTn algorithm (Altschul et al (1997) *Nucleic Acids Res.*

35 25:3389-3402.) against the Du Pont proprietary database comparing nucleotide sequences that share common or overlapping regions of sequence homology. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences can be assembled into a single contiguous nucleotide

sequence, thus extending the original fragment in either the 5 or 3 prime direction. Once the most 5-prime EST is identified, its complete sequence can be determined by Full Insert Sequencing. Homologous genes belonging to different species can be found by comparing the amino acid sequence of a known gene (from either a  
5 proprietary source or a public database) against an EST database using the tBLASTn algorithm. The tBLASTn algorithm searches an amino acid query against a nucleotide database that is translated in all 6 reading frames. This search allows for differences in nucleotide codon usage between different species, and for codon degeneracy.

10 cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene.  
15 Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into pre-cut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are  
20 prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed  
25 sequence tags or "ESTs"; see Adams et al., (1991) *Science* 252:1651-1656). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer. Full-insert sequence (FIS) data is generated utilizing a modified transposition protocol. Clones identified for FIS are recovered from archived glycerol stocks as single colonies, and plasmid DNAs are isolated via alkaline lysis.  
30 Isolated DNA templates are reacted with vector primed M13 forward and reverse oligonucleotides in a PCR-based sequencing reaction and loaded onto automated sequencers. Confirmation of clone identification is performed by sequence alignment to the original EST sequence from which the FIS request is made.

Confirmed templates are transposed via the Primer Island transposition kit  
35 (PE Applied Biosystems, Foster City, CA), which is based upon the *Saccharomyces cerevisiae* Ty1 transposable element (Devine and Boeke (1994) *Nucleic Acids Res.* 22:3765-3772). The *in vitro* transposition system places unique binding sites randomly throughout a population of large DNA molecules. The transposed DNA is

then used to transform DH10B electro-competent cells (Gibco BRL/Life Technologies, Rockville, MD) via electroporation. The transposable element contains an additional selectable marker (named DHFR; Fling and Richards (1983) *Nucleic Acids Res.* 11:5147-5158), allowing for dual selection on agar plates of only those subclones containing the integrated transposon. Multiple subclones are randomly selected from each transposition reaction, plasmid DNAs are prepared via alkaline lysis, and templates are sequenced (ABI Prism dye-terminator ReadyReaction mix) outward from the transposition event site, utilizing unique primers specific to the binding sites within the transposon.

Sequence data is collected (ABI Prism Collections) and assembled using Phred/Phrap (P. Green, University of Washington, Seattle). Phred/Phrap is a public domain software program which re-reads the ABI sequence data, re-calls the bases, assigns quality values, and writes the base calls and quality values into editable output files. The Phrap sequence assembly program uses these quality values to increase the accuracy of the assembled sequence contigs. Assemblies are viewed by the Consed sequence editor (D. Gordon, University of Washington, Seattle).

#### EXAMPLE 8

##### Characterization of cDNA Clones Encoding members of RTCS proteins

The BLASTX search using the sequences from clones listed in Table 1 below revealed similarity of the polypeptides encoded by the ORF to the hypothetical protein from rice (NCBI General Identifier No. 27261472, SEQ ID NO: 26) and a putative LOB domain protein from *Arabidopsis* [*Arabidopsis thaliana*] (NCBI General Identifier No. 15230971, which is identical to the amino acid sequence set forth in NCBI General Identifier No. 22331847 (SEQ ID NO:27)).

Table 3 shows the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), the sequences of contigs assembled from two or more ESTs ("Contig"), sequences of contigs assembled from an FIS and one or more ESTs ("Contig\*"), or sequences encoding an entire protein derived from an FIS, a contig, or an FIS and PCR ("CGS"):

TABLE 1

BLAST Results for Sequences Encoding the coding region of maize RTCS-like open reading frame and Polypeptides Homologous To RTCS

Sequence	Status	BLAST pLog Score	
		27261472	15230971
PCR product of RTCS gene from bac clone BAC74.m9	cgs	243	176

The data in Table 2 below represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:6, 8, 30, and 38 and the hypothetical protein from rice (NCBI General Identifier No. 27261472, SEQ ID NO:26) and the putative LOB domain protein from *Arabidopsis* [*Arabidopsis thaliana*] (NCBI General Identifier No. 22331847, SEQ ID NO:27).

TABLE 2

Percent Identity of Amino Acid Sequences encoded by Nucleotide Sequences of DNA sequences Encoding the RTCS protein and Polypeptides Homologous To RTCS

SEQ ID NO.	Percent Identity to	
	27261472	22331847
6	67.6	46.8
8	98.8	46.3
30	66.9	47.7
38	58.8	43.1

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant sequences encode a substantial portion of a LOB domain protein and share homology with the maize polypeptide

encoded by the maize *rtcs* gene, which gives rise to an altered root structure when mutated.

### EXAMPLE 9

#### Isolation of RTCS-LIKE Sequences

Leaves and roots were harvested off 10-day old corn B73 seedlings grown in liquid media consisting of half strength Hoagland salts, plus 7% sucrose added 24 hr prior to sampling. Total RNA was prepared from these samples after they were grounded in liquid nitrogen, using the Tri-PURE kit (Roche, Indianapolis, IN). Using the total RNA as templates, cDNA from the leaf and root samples was prepared using oligo(dT) primers with the Reverse Transcription System (Promega, Madison WI), following manufacturer's instruction.

For isolation of RTCS-LIKE sequences, two gene-specific primers were designed based on the coding sequences of the maize RTCS gene. PCR using the forward primer described in SEQ ID NO: 35, and the reverse primer described in SEQ ID NO:36, was performed in a PTC-200 DNA Engine thermal cycler from MJ Research Inc. (Waltham, Massachusetts) using reagents supplied with the ProofStart DNA Polymerase kit (Qiagen, Valencia, CA). The cDNA isolated from roots and leaves as described above served as a template for the reaction. The following cycle parameters were used: 5 min enzyme activation at 95°C, followed by 30 cycles of 94°C for 15 seconds, 55°C for 30 sec then 72°C for 1 minute. The samples were then held at 72°C for 10 minutes and then at 8°C until further analysis. Each PCR reaction mix was run on 1.0% agarose gel. Only root cDNA produced a PCR product. The PCR band was excised from gel, purified with the Qiaquick Gel Extraction Kit (Qiagen, Valencia, CA). The PCR product was then cloned into the TA cloning vector pCR2.1 (Invitrogen, Carlsbad, CA). The clones were sequenced for verification. The resulting cDNA clone encoding a RTCS-like gene is described in SEQ ID NO: 37. The protein encoded by SEQ ID NO: 37 is shown in SEQ ID NO: 38, and the genomic sequence containing the RTCS-like gene is shown in SEQ ID NO:39.

### EXAMPLE 10

#### Identification of a new mutant *rtcs* allele

A new mutant *rtcs* allele has been identified by screening a mu active population. DNA from *rtcs* plants was extracted and primers 296F (SEQ ID NO.: 40) and 1230R (SEQ ID NO.: 41) were used for PCR amplification. An unexpected PCR product size of 1536bp was obtained, indicating the presence of an insertion in the portion of the gene that was amplified. The mutant plants carry an insertion of 578bp



after nucleotide 831 of SEQ ID NO.: 28. Within the 578 additional bp a terminal repeat of 7bp (tcctgct) was found. The mutant plants also carry a A to G mutation at the splicing site of the intron at position 1573 of SEQ ID NO.: 28. The 3864 bp sequence containing the 1546 bp fragment with the mutant sequence is shown in SEQ ID NO.: 42 and Fig.4. Since the phenotype co-segregates perfectly with both the insertion and the mutation, RT-PCR experiments will be carried out in order to clarify whether both or only one of the above-described changes in the mutant can be held accountable for the *rtcs* phenotype.

### EXAMPLE 11

#### Expression of recombinant DNA Constructs in Monocot Cells

A recombinant DNA construct comprising a plant cDNA encoding the instant polypeptides in sense orientation with respect to promoter from the ubiquitin, or CaMV 35S, gene that is located 5' to the cDNA fragment can be constructed. The 3' fragment from the 10kD zein gene [Kiriara et al. (1988) *Gene* 71:359-370] can be placed 3' to the cDNA fragment. Such constructs are used to overexpress or cosuppress the gene(s) homologous to *RTCS*. It is realized that one skilled in the art could employ different promoters and/or 3'-end sequences to achieve comparable expression results. The construct with the CaMV 35S promoter is made as follows: the transcription termination element is released from the clone by digestion with appropriate restriction enzymes. The fragment is then ligated to appropriate restriction sites of pML141 [PCT Application No. WO 00/08162, published February 17, 2000], which carries the 35S promoter, using an appropriate linker. The DNA containing the *RTCS* ORF is amplified through PCR by using specific primer sets and the cDNA as a template. The fragment is then digested with appropriate restriction enzymes of the vector between the 35S promoter and the transcription terminator. The appropriate orientation of the insert is confirmed by sequencing.

The construct with the ubiquitin promoter is made as follows: the transcription termination element is released from the clone by digestion with the appropriate restriction enzymes digestion. The fragment is ligated to BamHI and NotI restriction sites of SK-ubi (BbsI), which carries the ubiquitin promoter (maize *Ubi-1* promoter, Christensen and Quail (1996) *Transgenic Res.* 5: 213-218), using the an appropriate linker. The DNA containing the *RTCS* ORF is amplified through PCR by using a specific primer set and the cDNA as a template. The fragment is then digested with appropriate restriction enzymes and inserted between the ubiquitin promoter and the transcription terminator.

Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb Sall-NcoI promoter fragment of the maize 27 kD zein gene [Prat et al. (1987) *Gene* 52:51-49; Gallardo et al. (1988) *PlantSci.* 54:211-2811] and a 0.96 kb SmaI-Sall fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene).

Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S. Biochemical). The resulting plasmid construct would comprise a recombinant DNA construct encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptides, and the 10 kD zein 3' region.

The recombinant DNA construct described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 µm in diameter) are coated with DNA using the following technique.

Ten  $\mu\text{g}$  of plasmid DNAs are added to 50  $\mu\text{L}$  of a suspension of gold particles (60 mg per mL). Calcium chloride (50  $\mu\text{L}$  of a 2.5 M solution) and spermidine free base (20  $\mu\text{L}$  of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200  $\mu\text{L}$  of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30  $\mu\text{L}$  of ethanol. An aliquot (5  $\mu\text{L}$ ) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains bialophos (5 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing bialophos. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the bialophos-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* 8:833-839).

## EXAMPLE 12

### Expression of Recombinant DNA Constructs in Dicot Cells

The 35S promoter of CaMV can be used to over-express and co-suppress the genes homologous to *RTCS* in dicot cells. For *RTCS* overexpression, the vector KS50 can be used to fuse the *RTCS* ORF to the 35S promoter. The *RTCS* ORF is amplified by PCR using a specific primer set. The amplified DNA fragment is digested with the appropriate restriction enzyme and ligated at the appropriate

site of KS50. The correct orientation of the insert is determined by sequencing. KS50 (7,453 bp) is a derivative of pKS18HH (U.S. Patent No. 5,846,784) which contains a T7 promoter/T7 terminator controlling the expression of a hygromycin phosphotransferase (HPT) gene, as well as a 35S promoter/NOS terminator  
5 controlling the expression of a second HPT gene. KS50 has an insert at the Sal I site consisting of a 35S promoter (960 bp)/NOS terminator (700 bp) cassette taken from pAW28, with a NotI cloning site between the promoter and terminator.

Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptides. To induce somatic  
10 embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos, which produce secondary embryos, are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos, which multiplied as early,  
15 globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

20 Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) 327:70-73, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean  
25 transformation is a recombinant construct composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al.(1983) *Gene* 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette  
30 comprising the phaseolin 5' region, the fragment encoding the instant polypeptides and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 µL of a 60 mg/mL 1 µm gold particle suspension is added (in order):  
35 5 µL DNA (1 µg/µL), 20 µL spermidine (0.1 M), and 50 µL CaCl<sub>2</sub> (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 µL 70% ethanol and resuspended in 40 µL of anhydrous

PL 1/0505/03332

ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five  $\mu\text{L}$  of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

### EXAMPLE 13

#### Identification of protein sequences specific to RTCS and RTCS homologs

LOB domain proteins comprise a family of genes involved in lateral organ boundary formation and all share a conserved amino acid sequence stretch of about 100 residues in the N-terminal regions of the proteins belonging to this family (Shuai et al. (2002) Plant Physiology : 129: 747-761). Figure 1 shows an alignment of the maize RTCS from inbred Mo17 (SEQ ID NO:6), a deduced RTCS homolog from rice (SEQ ID NO:8), maize RTCS from inbred B73 (SEQ ID NO:30), rice gi 27261472 (SEQ ID NO:26) and *Arabidopsis* gi: 22331847 (SEQ ID NO:27). The boxed residues are conserved motifs unique to RTCS proteins.